Probability Comprehensive Exam
Fall 2019

Student Number:

Instructions: Complete 5 of the 10 problems, and **circle** their numbers below – the uncircled problems will **not** be graded.

1 2 3 4 5 6 7 8 9 10

Write **only on the front side** of the solution pages. A student will pass the exam if 3 problems are worked "almost perfectly" and some progress is made on a fourth problem.

1. First question.

   Show that a random variable X such that
$$\mathbb{E}[e^{\lambda X}] \le e^{2|\lambda|^3} \text{ for any } \lambda \in [-1,1]$$
   satisfies $X = 0$ almost surely.

Solution: In the first place, if we apply this for λ and $-\lambda$ we get
$$\mathbb{E}[e^{|\lambda||X|}] \le \mathbb{E}[e^{\lambda X} + e^{-\lambda X}] \le 2e^{2|\lambda|^3}$$
Thus, X has all finite moments and in addition, we get that for $\lambda > 0$
$$\mathbb{E}[\frac{1}{\lambda}(e^{\lambda X} - 1)] \le \mathbb{E}[\frac{e^{2|\lambda|^3} - 1}{\lambda}]$$
and then from Fatou's Lemma, letting $\lambda \to 0$, we obtain that
$$\mathbb{E}[X] \le 0.$$
Replacing X by $-X$ we still have the inequality
$$\mathbb{E}[e^{-\lambda X}] \le e^{2|\lambda|^3}$$
and using the same argument as before, we obtain that
$$\mathbb{E}[-X] \le 0.$$
Combining the two, we arrive at $\mathbb{E}[X] = 0$. Next we use the second order expansion to get that
$$\mathbb{E}[\frac{1}{\lambda^2}(e^{\lambda X} - 1)] = \mathbb{E}[\frac{1}{\lambda^2}(e^{\lambda X} - 1 - \lambda X)] \le \frac{e^{2|\lambda|^3} - 1}{\lambda^2}$$
and now using again Fatou's lemma combined by letting $\lambda \to 0$, we get that
$$\mathbb{E}[X^2] = 0$$
which means $X = 0$ almost surely.

Another solution would be obtained by taking Y an independent copy of X and noticing that for $Z = X - Y$ we get that
$$\mathbb{E}[e^{\lambda Z}] \le e^{4|\lambda|^3}.$$
The advantage is that Z already has mean 0 and thus we can use the expansion to second order directly to get that
$$\mathbb{E}[Z^2] = 0$$
Now since this is the case, it means that $X = Y$ almost surely, which implies that X must be constant. Now going back to the inequality it is easy to see that the constant must be zero.

2. Assume $(X_n)_{n\geq 1}$ are iid (independent and identically distributed) random variables on some space $(\Omega, \mathcal{F}, \mathbb{P})$ with common Gumbel cumulative function given by

$$F(x) = e^{-e^{-x}}, x \in \mathbb{R}.$$

Show that

$$\limsup_{n\to\infty}(-X_n - \ln(\ln(n))) = 0.$$

Solution: We just have to prove two things.

$$1)\mathbb{P}(\limsup_{n\to\infty}(-\ln(\ln(n)) - X_n) > \delta) = 0$$
$$2)\mathbb{P}(\limsup_{n\to\infty}(-\ln(\ln(n)) - X_n) > -\delta) = 1.$$

To prove the first part we use the first implication of Borel-Cantelli to justify

$$\begin{aligned}\sum_{n\geq 1}\mathbb{P}(-\ln(\ln(n)) - X_n > \delta/2) &= \sum_{n\geq 1}\mathbb{P}(X_n \leq -\ln(\ln(n)) - \delta/2) \\ &= \sum_{n\geq 1} e^{-e^{\ln(\ln(n))+\delta/2}} = \sum_{n\geq 1}\frac{1}{n^{e^{\delta/2}}} < \infty\end{aligned}$$

and similarly

$$\begin{aligned}\sum_{n\geq 1}\mathbb{P}(-\ln(\ln(n)) - X_n > -\delta/2) = \sum_{n\geq 1}\mathbb{P}(X_n \leq -\ln(\ln(n)) + \delta/2) = \sum_{n\geq 1} e^{-e^{\ln(\ln(n))-\delta/2}} \\ = \sum_{n\geq 1}\frac{1}{n^{e^{-\delta/2}}} = \infty.\end{aligned}$$

Thus the Borel-Cantelli does the trick.

The interpretation is that the $-X_n - \ln(\ln(n))$ does not grow more than 0 in the long run and it stays infinitely many times around close to 0.

A nicer interpretation is that

$$\limsup_{n\to\infty}(-X_n - \ln(\ln(n))) = -\liminf_{n\to\infty}(X_n + \ln(\ln(n)))$$

so,

$$\liminf_{n\to\infty}(X_n + \ln(\ln(n))) = 0$$

almost surely. This is actually easier to interpret, because it means for any $\delta > 0$, after a while, $X_n + \ln(\ln(n)) > -\delta$. In other words, for large enough n's, roughly, we have that $X_n \geq -\ln(\ln(n))$.

3. Let $(a_i)_{i\geq 1}$ be a sequence of positive integers such that $a_i \in [\lceil 1.01^{i-1}\rceil, \lceil 1.01^{i}\rceil]$ for all $i \geq 1$. Further, let $(S_n)_{n\geq 0}$ be a random walk on $\mathbb{Z}$, with $S_0 = 0$ and with $S_n = \sum_{i=1}^n X_i$, where $(X_i)_{i\geq 1}$ are mutually independent random variables, with $\mathbb{E}[X_i] = 0$ for all $i \geq 1$, and $|X_i| = a_i$, $i \geq 1$, everywhere on the probability space. Prove that the random walk (S_n) is not recurrent.

Solution: To show that the random walk is not recurrent, it is sufficient to verify that

$$\sum_{n=1}^{\infty} \mathbb{P}\{S_n = 0\} < \infty.$$

We will prove a much stronger inequality; namely, that $\big(\mathbb{P}\{S_n = 0\}\big)_{n\geq 1}$ is majorized by a geometric progression.

First, we observe that by the definition of (a_i) there is a universal integer constant $C > 0$ such that for any sequence of signs $(\varepsilon_j)_{j\geq 1}$ we have

$$\sum_{i=1}^{r} \varepsilon_i a_i + \sum_{i=r+1}^{r+C} a_i > 0 \quad \text{for all } r \geq 1.$$

For convenience, let us agree that $\operatorname{sign}(0) = 1$ below. Now, take any $n \geq 2C$, set $m := \lfloor n/C \rfloor$ and observe that, by the above condition, the event $\{S_n = 0\}$ must be contained inside the event

$$\Big\{\operatorname{sign}(X_i) \neq \operatorname{sign}(X_j) \text{ for at least one pair } i,j \in [n-C+1, n] \ \text{ AND}$$

$$\operatorname{sign}(X_i) \neq \operatorname{sign} \sum_{q=n-C+1}^{n} X_q \text{ for at least one index } i \in [n-2C+1, n-C] \ \text{ AND}$$

$$\dots$$

$$\operatorname{sign}(X_i) \neq \operatorname{sign} \sum_{q=n-(m-1)C+1}^{n} X_q \text{ for at least one index } i \in [n-mC+1, n-(m-1)C]\Big\}.$$

By the independence, probability of the last event can be estimated from above by $\big(1 - 2^{-C}\big)^m$. The result follows.

4. 1. If X is a random variable such that for two constants $a, b \in \mathbb{R}$, we have $a \leq X \leq b$, show that $\text{var}(X) \leq (b-a)^2/4$ and give an example of such a random variable where equality is attained.

   2. Assume that X is a random variable such that $\mathbb{P}(X \leq a) = 1/2$ and $\mathbb{P}(X \geq b) = 1/2$ for some real numbers a, b, $a < b$. Show that $\text{var}(X) \geq (b-a)^2/4$ and give an example of such a random variable where equality is attained.

Solution:

1. There are two ways of proving this. The first one is using the following simple characterization of the variance as

$$\text{var}(X) = \min_{t \in \mathbb{R}} \mathbb{E}[(X-t)^2].$$

For our case, taking $t = (a+b)/2$ and using the fact that $|X - \frac{a+b}{2}| \leq (b-a)/2$, we immediately get the conclusion.

The other way of doing it is to reduce it to the case of $a = -1$, $b = 1$ by linear scaling. For instance, defining $Y = 2(X - (a+b)/2)/(b-a)$, we obviously have that $\text{var}(Y) = \frac{4}{(b-a)^2}\text{var}(X)$. Since $|Y| \leq 1$, we have that

$$\text{var}(Y) = \mathbb{E}[Y^2] - \mathbb{E}[Y]^2 \leq 1$$

and this transfers to $\text{var}(X) \leq \frac{(b-a)^2}{4}$.

One can trace both cases of equality and realize that this is attained for the case of $\mathbb{P}(X = a) = \mathbb{P}(X = b) = 1/2$. Perhaps this is most transparent through the second solution because in that case we can plainly see that $\text{var}(Y) = 1$ if and only if $Y^2 = 1$ and $\mathbb{E}[Y] = 0$ which basically means that $Y = \pm 1$, each value being equally likely.

2. In a similar vein to the first part we can treat this item too. Again since the variance is

$$\text{var}(X) = \min_{t \in \mathbb{R}} \mathbb{E}[(X-t)^2].$$

Now, we are going ot split the analysis according to $t \leq a$, $t \geq b$ and $a \leq t \leq b$. For $t \leq a$, we have

$$\begin{aligned}\mathbb{E}[(X-t)^2] &= \mathbb{E}[(X-t)^2, X \leq a] + \mathbb{E}[(X-t)^2, X \geq b] \\ &\geq \mathbb{E}[(X-t)^2, X \geq b] \geq (b-a)^2 \mathbb{P}(X \geq b) = \frac{(b-a)^2}{2}\end{aligned}$$

In a similar fashion we get that for $t \geq b$, we get $\mathbb{E}[(X-t)^2] \geq \frac{(b-a)^2}{2}$.

For $a \le t \le b$, we have

$$\begin{aligned}\mathbb{E}[(X-t)^2] &= \mathbb{E}[(X-t)^2, X \le a] + \mathbb{E}[(X-t)^2, X \ge b] \\ &\ge (t-a)^2\mathbb{P}(X \le a) + (b-t)^2\mathbb{P}(X \ge b) \\ &= \frac{1}{2}(2t^2 - 2t(a+b) + a^2 + b^2) \\ &= \left(t - \frac{a+b}{2}\right)^2 + \frac{(b-a)^2}{4} \ge \frac{(b-a)^2}{4}.\end{aligned}$$

It is now clear that $\mathrm{var}(X) \ge \frac{(b-a)^2}{4}$ and equality is achieved for the case of $p(X = a) = 1/2$ and $\mathbb{P}(X = b) = 1/2$.

Another solution is based on the the fact that if we take X' another independent copy of X, then

$$\begin{aligned}\mathrm{var}(X) = \frac{1}{2}\mathbb{E}[(X-X')^2] = \frac{1}{2}\Big(&\mathbb{E}[(X-X')^2, X, X' \le a] + \mathbb{E}[(X-X')^2, X, X' \ge b] \\ &+ 2\mathbb{E}[(X-X')^2, X \le a, X' \ge b]\Big) \ge \mathbb{E}[(X-X')^2, X \le a, X' \ge b] \\ &\ge (b-a)^2\mathbb{P}(X \le a, X' \ge b) = \frac{(b-a)^2}{4}.\end{aligned}$$

Equality is attained if X is Bernoulli with $\mathbb{P}(X = a) = \mathbb{P}(X = b) = 1/2$.

5. Let $(S_n)_{n\geq 1}$ be a simple random walk on $\mathbb{Z}$ (with $S_0 = 0$). Compute the probability mass function of the maximum of the random walk on the interval $[2n, 4n]$, i.e. compute the pmf of the variable $\xi := \max_{2n\leq i\leq 4n} S_i$. Represent the pmf as a (polynomial) function of binomial coefficients.

Solution: Take any integer $m > -2n$. Applying the reflection principle, we get

$$\begin{aligned}\mathbb{P}\big\{\xi \geq m\big\} &= \mathbb{P}\big\{S_{2n} \geq m\big\} + \sum_{i=-2n}^{m-1} \mathbb{P}\big\{S_{2n} = i \text{ and } \xi - S_{2n} \geq m-i\big\} \\ &= \mathbb{P}\big\{S_{2n} \geq m\big\} + \sum_{i=-2n}^{m-1} \mathbb{P}\big\{S_{2n} = i\big\} \big(\mathbb{P}\{S_{2n} = m-i\} + 2\mathbb{P}\big\{S_{2n} > m-i\big\}\big).\end{aligned}$$

This implies that, taking any $k > -n$, we have

$$\begin{aligned}\mathbb{P}\big\{\xi = 2k\big\} &= \mathbb{P}\{\xi \geq 2k\} - \mathbb{P}\{\xi \geq 2k+2\} \\ &= \mathbb{P}\big\{S_{2n} = 2k\big\} + \sum_{i=-2n}^{2k-1} \mathbb{P}\big\{S_{2n} = i\big\} \big(\mathbb{P}\{S_{2n} = 2k-i\} + 2\mathbb{P}\big\{S_{2n} > 2k-i\big\}\big) \\ &\qquad - \sum_{i=-2n}^{2k+1} \mathbb{P}\big\{S_{2n} = i\big\} \big(\mathbb{P}\{S_{2n} = 2k+2-i\} + 2\mathbb{P}\big\{S_{2n} > 2k+2-i\big\}\big) \\ &= \mathbb{P}\big\{S_{2n} = 2k\big\} + \sum_{i=-n}^{k-1} \mathbb{P}\big\{S_{2n} = 2i\big\} \big(\mathbb{P}\{S_{2n} = 2k-2i\} + 2\mathbb{P}\big\{S_{2n} > 2k-2i\big\}\big) \\ &\qquad - \sum_{i=-n}^{k} \mathbb{P}\big\{S_{2n} = 2i\big\} \big(\mathbb{P}\{S_{2n} = 2k+2-2i\} + 2\mathbb{P}\big\{S_{2n} > 2k+2-2i\big\}\big) \\ &= \mathbb{P}\big\{S_{2n} = 2k\big\} - \mathbb{P}\big\{S_{2n} = 2k\big\} \big(\mathbb{P}\{S_{2n} = 2\} + 2\mathbb{P}\big\{S_{2n} > 2\big\}\big) \\ &\quad + \sum_{i=-n}^{k-1} \mathbb{P}\big\{S_{2n} = 2i\big\} \big(\mathbb{P}\{S_{2n} = 2k-2i\} + \mathbb{P}\big\{S_{2n} = 2k+2-2i\big\}\big) \\ &= \mathbb{P}\big\{S_{2n} = 2k\big\}\big(\mathbb{P}\{S_{2n} = 2\} + \mathbb{P}\{S_{2n} = 0\}\big) \\ &\quad + \sum_{i=-n}^{k-1} \mathbb{P}\big\{S_{2n} = 2i\big\} \big(\mathbb{P}\{S_{2n} = 2k-2i\} + \mathbb{P}\big\{S_{2n} = 2k+2-2i\big\}\big) \\ &= \sum_{i=-n}^{k} \mathbb{P}\big\{S_{2n} = 2i\big\} \big(\mathbb{P}\{S_{2n} = 2k-2i\} + \mathbb{P}\big\{S_{2n} = 2k+2-2i\big\}\big).\end{aligned}$$

Switching to the binomial coefficients, we get

$$\mathbb{P}\big\{\xi = 2k\big\} = \sum_{i=-n}^{k} \binom{2n}{n-i} \left(\binom{2n}{n-k+i} + \binom{2n}{n-k+i-1}\right).$$

6. Assume $(X_n)_{n\geq 1}$ is a sequence of iid positive random variables. Show that

$$\frac{X_1 + X_2^2 + \cdots + X_n^n}{n} \xrightarrow[n\to\infty]{a.s.} 1 \text{ if and only if } X_1 = 1 \text{ almost surely.}$$

Solution: There are two parts, the first is to show that $\mathbb{P}(X_1 > 1+\epsilon) = 0$ and the second part that $\mathbb{P}(X_1 < 1-\epsilon) = 0$.

For the first part, fix an integer value k and use that

$$\begin{aligned}\frac{X_1 + X_2^2 + \cdots + X_n^n}{n} &\geq \frac{(1+\epsilon)\mathbb{1}_{[1+\epsilon,\infty)}(X_1) + (1+\epsilon)^2\mathbb{1}_{[1+\epsilon,\infty)}(X_2) + \cdots + (1+\epsilon)^n\mathbb{1}_{[1+\epsilon,\infty)}(X_n)}{n} \\ &\geq (1+\epsilon)^k \frac{\mathbb{1}_{[1+\epsilon,\infty)}(X_k) + \mathbb{1}_{[1+\epsilon,\infty)}(X_k) + \cdots + \mathbb{1}_{[1+\epsilon,\infty)}(X_n)}{n}\end{aligned}$$

Consequently, letting $n \to \infty$, by the law of large numbers and the assumption of the problem, we obtain that

$$1 \geq (1+\epsilon)^k \mathbb{P}(X_1 \geq 1+\epsilon)$$

for any $\epsilon > 0$ and any k. This is impossible unless $\mathbb{P}(\mathbb{P}(X_1 \geq 1+\epsilon)) = 0$ for any $\epsilon > 0$. Consequently, $X_1 \leq 1$ almost surely.

For the second part, again take $0 < \epsilon < 1$ and use now that almost surely we have

$$\begin{aligned}\frac{X_1 + X_2^2 + \cdots + X_n^n}{n} &\leq \frac{(1-\epsilon)\mathbb{1}_{[0,1-\epsilon)}(X_1) + (1-\epsilon)^2\mathbb{1}_{[0,1-\epsilon)}(X_2) + \cdots + (1-\epsilon)^n\mathbb{1}_{[0,1-\epsilon)}(X_n)}{n} \\ &\quad + \frac{\mathbb{1}_{[1-\epsilon,1]}(X_1) + \mathbb{1}_{[1-\epsilon,1]}(X_2) + \cdots + \mathbb{1}_{[1-\epsilon,1]}(X_n)}{n} \\ &\leq \frac{(1-\epsilon) + (1-\epsilon)^2 + \cdots + (1-\epsilon)^n}{n} \\ &\quad + \frac{\mathbb{1}_{[1-\epsilon,1]}(X_1) + \mathbb{1}_{[1-\epsilon,1]}(X_2) + \cdots + \mathbb{1}_{[1-\epsilon,1]}(X_n)}{n}\end{aligned}$$

Again, by letting $n \to \infty$ and using the law of large numbers we get that $\mathbb{P}(X_1 \geq 1-\epsilon) \geq 1$, for any $\epsilon > 0$, which means that $\mathbb{P}(X_1 = 1) = 1$.

7. (Modified Polya's urn) Consider the following discrete time process. Before time one, we have one white and one black ball in the urn. At time k ($k \geq 1$), we pick a ball from the urn uniformly at random **with replacement**, and add to the urn a ball of the color opposite to the color of the ball we have picked. Thus, at each step the number of balls in the urn increases by one. Let X_n be the proportion of white balls in the urn right after the n-th step. Show that (X_n) converges to $1/2$ almost surely.

Solution: By the symmetry, it is sufficient to show that for any $\varepsilon > 0$ we have

$$\mathbb{P}\{X_m \geq 1/2 + \varepsilon \text{ for some } m \geq n\} \to 0$$

as n goes to infinity. We will prove a stronger statement; namely that the sequence $\big(\mathbb{P}\{X_n \geq 1/2 + \varepsilon\}\big)$ is majorized by a (convergent) geometric sequence.

Fix any (large) n and any $1 \leq \ell \leq n+1$ such that $\ell \geq (1/2+\varepsilon)(n+2)$. We will estimate the probability of the event $\mathcal{E} := \{(n+2)X_n = \ell\}$, i.e. that after n steps we have ℓ white and $n+2-\ell$ black balls in the urn. This event means that in the course of the game we chose a black ball $\ell - 1$ times, and a black ball — $n - \ell + 1$ times. The event is then split into $\binom{n}{\ell-1}$ mutually disjoint subevents of the form $\mathcal{E}_{(c_1 c_2 \dots c_n)}$, where c_i can take values "b" or "w" and denotes the color of the ball picked at i-th step (so that there are $\ell - 1$ indices i with $c_i =$"b").

Instead of estimating probabilities of $\mathcal{E}_{(c_1 c_2 \dots c_n)}$ for any admissible sequence, we will compare probabilities of those events to identify the events with largest probabilities. Set

$$c := (c_1 c_2 \dots c_n), \quad c' := (c_1 c_2 \dots c_{i-1} c_{i+1} c_i c_{i+2} \dots c_n),$$

i.e. sequence c' is obtained from c by switching c_i and c_{i+1}. Assume that $c_i =$"b", $c_{i+1} =$"w", and that u is the number of white balls in the urn before the i-th step. Then the probabilities of $\mathcal{E}_c$ and $\mathcal{E}_{c'}$ are related by

$$\frac{\mathbb{P}(\mathcal{E}_{c'})}{\mathbb{P}(\mathcal{E}_c)} = \frac{u/(i+1) \,\cdot\, (i+2-u)/(i+2)}{(i+1-u)/(i+1) \,\cdot\, (u+1)/(i+2)} = \frac{ui - u^2 + 2u}{ui - u^2 + i + 1}.$$

Thus, whenever $i < 2u$, the probability of $\mathcal{E}_{c'}$ is not less than $\mathbb{P}(\mathcal{E}_c)$. Applying this observation, we get that a sequence $\widetilde{c}$ corresponding to $\mathcal{E}_{\widetilde{c}}$ of largest possible probability, has the form $\{b,w\}\{b,w\}\{b,w\}bb\dots b$, where $\{b,w\}$ means that the order of the pair bw or wb is not important (corresponds to the same probability). This gives

$$\mathbb{P}(\mathcal{E}_c) \leq \frac{1}{2} \cdot \frac{2}{3} \cdot \frac{1}{2} \cdot \frac{3}{5} \cdot \frac{1}{2} \cdot \frac{4}{7} \cdot \dots \cdot \frac{n+2-\ell}{2n+4-2\ell} \cdot \frac{n+2-\ell}{2n+5-2\ell} \cdot \frac{n+2-\ell}{2n+6-2\ell} \cdot \dots \cdot \frac{n+2-\ell}{n+1}.$$

It can then be shown that there is $C_\varepsilon > 0$ depending only on ε, and a universal constant $c > 0$ such that

$$\mathbb{P}(\mathcal{E}_c) \leq C_\varepsilon \left(\frac{1}{2} - c\varepsilon^2\right)^n.$$

Taking the union over 2^n admissible sequences c, we get that $\mathbb{P}(\mathcal{E})$ is exponentially small in n. The result follows.

8. Let $b_1, b_2, \ldots$ be a sequence of mutually independent Bernoulli(1/2) random variables, and let m be a fixed positive integer. Define a process $(X_i)_{i\geq 0}$ as follows. Set $X_0 = 1$, and define X_i recursively as

$$X_i = \begin{cases} X_{i-1} + \frac{1}{m}(2b_i - 1), & \text{if } X_{i-1} \geq 1/m; \\ 0, & \text{otherwise.} \end{cases}$$

Show that (X_n) converges to zero almost everywhere.

Solution: By the definition, (X_n) is a non-negative martingale. Indeed, it is clear that X_n is a non-negative process. It is a martingale because

$$\mathbb{E}[X_{n+1}|\mathcal{F}_n] = X_n \mathbb{1}_{X_n \geq 1/m} = X_n$$

The last equality comes from the fact that X_n takes values in $\mathbb{Z}/m$, thus $X_n \mathbb{1}_{X_n \geq 1/m} = X_n \mathbb{1}_{X_n > 0} = X_n$. Without this extra bit (that X_n takes values in $\mathbb{Z}/m$) we would have gotten anyhow that

$$\mathbb{E}[X_{n+1}|\mathcal{F}_n] = X_n \mathbb{1}_{X_n \geq 1/m} \leq X_n$$

which means that X_n is a non-negative supermatingale.

Then, by the Martingale Convergence Theorem, there is a random variable X such that X_n converges to $X \geq 0$ almost everywhere. To prove that $X = 0$ a.e. observe that almost everywhere on the event $\{X \neq 0\}$ we have $\sum_{i=1}^{k}(2b_i - 1) > -m$ for all $k \geq 1$. This is equivalent to saying that the minimum of a simple random walk on $\mathbb{Z}$ is bounded from below by $-m+1$. But the latter has probability zero, hence the result.

9. Let $m \geq 2$ be a positive integer, and let $b_1, b_2, \ldots$ be mutually independent Bernoulli(1/2) variables. Consider the following Markov chain (X^n) in $\mathbb{R}^m$. Let X^0 be a fixed 0/1–vector in $\mathbb{R}^m$. Next, given $X^{i-1} = (x_1^{i-1}, x_2^{i-1}, \ldots, x_m^{i-1})$, we set $X^i = (x_1^i, x_2^i, \ldots, x_m^i)$ to be the 0/1–vector such that

$$\sum_{j=1}^{m} 2^{n-j} x_j^i - \sum_{j=1}^{m} 2^{n-j} x_j^{i-1} = \begin{cases} b_i(1-2^m), & \text{if } x_1^{i-1} = x_2^{i-1} = \cdots = x_m^{i-1} = 1; \\ b_i, & \text{otherwise.} \end{cases}$$

(Above, "$i-1$", "i" are upper indices, not powers)

1) Prove that the Markov chain (X^n) converges in distribution to the uniform distribution on the set $\{0,1\}^m$.

2) Recall that the mixing time t_{mix} is defined as the smallest integer such that for all $n \geq t_{mix}$ the total variation distance between the distribution of X^n and the stationary distribution is at most $1/4$. Show that the mixing time t_{mix} of (X^n) satisfies $c\,2^{2m} \leq t_{mix} \leq C\,2^{2m}$ for some universal constants $c, C > 0$. You should not use as a blackbox "known" estimates on mixing times, please outline the proof.

Solution: First of all, it is convenient to view the chain as a lazy walk on the set of numbers from 0 to $2^m - 1$ (written in binary notation), with the convention that the walk proceeds from number $2^m - 1$ to 0 (with probability 1/2); or as a lazy walk on a cycle of length 2^m.

The first part of the problem follows immediately from the convergence theorem for finite state Markov chains: because of the laziness, the chain is aperiodic; it is also clearly irreducible.

For the lower bound in the second part, assume for concreteness that the walk starts with 0. We can assume without loss of generality that m is large. After $c\,2^{2m}$ steps, the position of the walk is $\sum_{i=1}^{c\,2^{2m}} b_i \mod 2^m$. It is easy to see that the variance of $\xi := \sum_{i=1}^{c\,2^{2m}} b_i$ is equal to $\frac{c}{4}\,2^{2m}$. Hence, taking a sufficiently small c, we get by Markov's inequality that with probability, say, at least 0.9 the variable ξ falls into an interval of length $2^m/2$. Then the measure of corresponding collection of states (induced by the distribution of $X^{c\,2^{2m}}$) is at least 0.9 while the stationary measure of the collection is 1/2. Thus, the total variation distance is greater than 1/4, implying the lower bound.

For the upper bound, we will compare probabilities of events $\{X^n = s_1\}$ and $\{X^n = s_2\}$ (for $n \geq C\,2^{2m}$ for a large universal constant C), for any two states s_1 and s_2.

We have

$$\mathbb{P}\big\{X^n = s_1\big\} = \mathbb{P}\Big\{\sum_{i=1}^n b_i \equiv s_1 \mod 2^m\Big\} = 2^{-n} \sum_{i:\, i\equiv s_1 \mod 2^m} \binom{n}{i};$$
$$\mathbb{P}\big\{X^n = s_2\big\} = 2^{-n} \sum_{i:\, i\equiv s_2 \mod 2^m} \binom{n}{i}.$$

Standard computations imply that for any $c_1 > 0$ there is C depending only on c_1 such that for all $n \geq C\, 2^{2m}$ we have

$$\left|\binom{n}{i} - \binom{n}{j}\right| \leq \min\left(e^{-c'(i-n/2)^2/n}, c_1\right)\frac{2^n}{\sqrt{n}},$$

whenever $|i-j| \leq 2^m$, where $c' > 0$ is a universal constant. Using this estimate, we get that as long as C is sufficiently large, we have

$$\frac{\mathbb{P}\{X^n = s_1\}}{\mathbb{P}\{X^n = s_2\}} \in [0.99, 1.01].$$

This immediately implies that the total variation distance to the uniform distribution on $\{0,1\}^m$ is less than $1/4$, hence the result.

10. On a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ assume we have three random variables X, Y, Z independent and uniform on $[0,1]$. Compute $\mathbb{E}[\min\{X, Y, Z\}]$.

Solution:

The idea is to first compute it's distribution function of $U = \min\{X, Y, Z\}$. This is not so complicates as for $0 \leq u \leq 1$, we have

$$F_U(u) = \mathbb{P}(\min\{X,Y,Z\} \leq u) = 1 - \mathbb{P}(\min\{X,Y,Z\} > u) = 1 - \mathbb{P}(X > u, Y > u, Z > u) = 1 - (1-u)^3$$

obviously for $u < 0$, $F_U(u) = 0$ while for $u > 1$, $F_U(u) = 1$. Now the density of U is

$$f_U(u) = \begin{cases} 3(1-u)^2 & 0 < u < 1 \\ 0 & \textit{otherwise.} \end{cases}$$

from which we immediately have

$$\mathbb{E}[U] = \int_0^1 u f_U(u) du = \int_0^1 3u(1-u)^2 du = 1/4.$$

An alternative argument is based on the splitting

$$\mathbb{E}[U] = \mathbb{E}[X, X < Y, Z] + \mathbb{E}[Y, Y < X, Z] + \mathbb{E}[Z, Z < X, Y].$$

By symmetry it is enough to compute one of these, the other two being the same.

$$\mathbb{E}[X, X < Y, Z] = \iiint_{0 \leq x < y \leq 1, 0 < x < z \leq 1} x dx dy dz = \int_0^1 \left(\int_x^1 \int_x^1 dy dz \right) x dx = \int_0^1 x(1-x)^2 dx = 1/12.$$

which leads to the same answer as above.